GZA Ziekenhuizen GasthuisZusters Antwerpen Sint-Augustinus - Sint-Vincentius - Sint-Jozef

Onze-Lieve-Vrouwziekenhuis Aalst - Asse - Ninove

CAT

Critically Appraised Topic

Optimizing an internal quality control program of automated ANA IIF

Auteur: Laura Bogaert

Supervisor: Apr. Sofie Schouwers, Apr. Lieve Van Hoovels

Datum: 08/05/2018

CLINICAL BOTTOM LINE

The detection of anti-nuclear antibodies (ANA) is important for the diagnosis of antibody associated rheumatic diseases (AARDs). The gold standard screening method for ANA, indirect immunofluorescence assay (IIF) on human epithelial (HEp-2) cells, is burdened with a number of disadvantages like a high workload, subjective visual reading and a consequently high intra- and inter-laboratory variance. The introduction of automated microscopic analysis may allow for more harmonized ANA IIF reporting, provided that a thorough quality assurance program controls this fully automated process. This program implies a continuous internal quality control that covers the total ANA IIF process, from pre- to post-analytical phase.

The aim of our study is to evaluate quality indicators used for ANA IIF analysis and a cost-effective optimization of the internal quality control program, covering the total ANA IIF analytical process.

CLINICAL/DIAGNOSTIC SCENARIO

Anti-nuclear antibodies (ANA) play a key role in the diagnosis, classification and prognosis of antibody associated rheumatic diseases (AARDs), such as systemic lupus erythematosus, systemic sclerosis, Sjögren's syndrome and idiopathic inflammatory myopathies (1, 2). According to recent recommendations of the American College of Rheumatology, ANA Task Force, the indirect immunofluorescence assay (IIF) on human epithelial (HEp-2) cells remains the gold standard for ANA testing (3). ANA IIF on HEp-2 cells, showing a multitude of native antigens, can be considered as a 'natural array' multiplex technique that allows the detection of > 30 different nuclear and cytoplasmic antigens and corresponding patterns (4, 5). However, conventional ANA IIF testing is time-consuming, laborious, and burdened by the need for microscopy expertise, subjectivity of interpretation, lack of automated procedures and the high variability of cellular substrates directly implying high intra- and inter-laboratory variance (6, 7). As a result of these limitations and the growing request of autoimmune diagnostic tests (8), automated ANA IIF systems have been introduced in autoimmunity laboratories (9). Nowadays, there are different commercial systems available (10-15). These systems differ in terms of DNA counterstain, software algorithms for IIF detection and pattern recognition, run-time, types of recognized ANA IIF patterns and their ability to analyze different substrates. Despite these differences, scientific literature suggests that these systems may contribute to the harmonization of the HEp-2 IIF analysis (7, 16) This was confirmed by a Belgian multicenter

study among laboratories performing ANA detection by NOVA View (Inova Diagnostics, San Diego, USA), demonstrating a lower ANA IIF titer variability in comparison to manual ANA IIF. However, for two laboratories clinically important deviations were found, due to pre-analytical and analytical problems, not revealed by their quality control scheme, solely based on company internal quality control (iQC) materials (17). A more recent Belgian inter-laboratory survey, evaluating variation in ANA detection by different automated IIF systems [NOVA View, EUROPattern (Euroimmun, Lübeck, Germany); G-Sight (Menarini, Firenze, Italy); Image Navigator (ImmunoConcepts, Sacramento, California)], not only found variation between automated IIF analysis using instruments from different manufacturers but also between instruments from the same manufacturer (18). Efforts should be undertaken to harmonize automated IIF analysis (16, 18). This could include the use of standards, calibration of the instruments and monitoring of the quality of the slides and reagents. The latter implies the introduction of a continuous iQC scheme that covers the total ANA IIF process, from pre- to post-analytical phase (17). Automated ANA IIF systems in the autoimmunity laboratory enable the introduction of objective iQC procedures to monitor the total ANA IIF process and prevent clinical significant shifts of ANA IIF measurements (19, 20). To establish a thorough quality assurance system, well-defined quality indicators and iQC performance criteria are needed. The final purpose of this study is to optimize the iQC program of the QUANTA-Lyser-NOVA View system.

QUESTION(S)

1) Defining usable quality indicators to reveal analytical and clinically significant errors in the total, automated ANA IIF process.
2) Evaluation of different quality indicators in an experimental setup.
3) Evaluation of different quality indicators in daily routine laboratory practice.

SEARCH TERMS

1) MeSH Database (PubMed): MeSH term: "antinuclear antibody"
2) PubMed Clinical Queries (from 1966; http://www.ncbi.nlm.nih.gov/entrez/query.fcgi): Systematic Reviews; Clinical Queries using Research Methodology Filters ("antinuclear antibody and automation", "antinuclear antibody and standardization", "antinuclear antibody and quality control system", "indirect immunofluorescence and antinuclear antibody and quality assurance")
3) Pubmed (Medline; from 1966), "antinuclear antibody and standardization", 'autoantibodies and harmonization", "indirect immunofluorescence and antinuclear antibody and quality assurance".
4) National Committee for Clinical Laboratory Standards (NCCLS; http://www.nccls.org/), Westgard QC (http://www.westgard.com)
5) UpToDate Online version (2016)

1) Defining usable quality indicators to reveal analytical and clinical significant errors in the total, automated ANA IIF process.

Materials and methods

QUANTA-Lyser – NOVA View system

The QUANTA-Lyser 2 instrument (Inova Diagnostics, San Diego, USA) is a pre-analytical platform for ANA IIF which automatically performs sample dilution and HEp-2 slide (NOVA Lite HEp-2 ANA kit, Inova Diagnostics, San Diego, USA) processing according to the instructions of the manufacturer.

NOVA View is an automated fluorescent microscope programmed to acquire, archive and manage digital images of fluorescent stained slides. The system encloses an Olympus 1x81 inverted IIF microscope with 4x, 10x and 40x objectives and dual band DAPI/ FITC/HC filters, computer, LED light source and a Kappa DX4 digital camera. The LED UV light source is a CoolLed PreciseExcite with excitation wavelengths of 400 nm (DAPI) and 490 nm (FITC). DAPI fluorescence is used by the NOVA View software for localizing the HEp-2 cells and focusing. Thereafter, the image analysis is performed based on the FITC signal. For each well in a slide, depending on the laboratory specific settings, three to eight images are acquired with both the DAPI and the FITC filter. Using FITC images, the system measures the average fluorescence intensity (FI) in units, i.e. light intensity units (LIU), discriminating between positive and negative samples. The cut-off set by Inova for ANA IIF positivity is 48 LIU. The NOVA View is able to identify and suggest five basic fluorescent ANA patterns (homogeneous, speckled, centromere, nucleolar and nuclear dots) based on software algorithms. Using pattern-specific dilution curves, the measured LIU can be converted into an estimated endpoint titer (single well titer (SWT)) (17, 21).

Quality materials

iQC materials are fit for purpose if, on the one hand, they represent the whole ANA IIF analytical process, from dilution up to result interpretation and on the other hand their variability uncovers the most important process errors. Positive ANA IIF serum samples with FI's corresponding to low titer (1:160) ANA IIF positivity, together with negative ANA IIF serum samples showing a low FI, reveal the most information regarding ANA IIF quality assurance (17).

Therefore, in addition to the company iQC materials NOVA Lite HEp-2 ANA kit, we have selected routine patient samples as iQC material. For the negative sample iQC, anonymized ANA IIF negative rest samples were pooled. For the positive sample iQC, two different pools were made, one with a speckled ANA IIF pattern and one with a homogeneous pattern, both targeting a titer of 200 LIU, resulting in a corresponding SWT of 1:160.

Quality indicators

The availability of a quantitative measure for FI (LIU), generated by automated ANA IIF microscopes, enables the use of IQC procedures generally applied in the automated chemistry laboratory (19, 20). Quality assurance can rely on the daily follow-up of LIU values of positive and negative iQC measurements based on the traditionally used Westgard multirules (22, 23). Additionally, it is worthwhile to include quality markers for the whole ANA

IIF testing process in the daily routine iQC analysis, for example the median patient LIU for every routine run (19). Such markers are independent on the sample position used and have shown to be of added value in the quality management of automated ANA IIF, when initial lot-to-lot comparison protocols fail to detect a change. Based on those principles, we have decided on the evaluation of different parameters (Table I) as a useful quality indicator for the ANA IIF process. For the determination of the median, mean and percentage positive ANA IIF patient samples LIU for every run, only results obtained for the 1:80 screening dilution were included.

Table 1. Quality indicators

Quality indicators
LIU positive kit iQC
LIU negative kit iQC
LIU positive sample iQC (Speckled)
LIU positive sample iQC (Homogeneous)
LIU negative sample iQC
% positive ANA IIF patient samples/run
Median patient sample LIU/run
Mean patient sample LIU/run

iQC acceptance criteria

To make objective decisions regarding analytical and clinical differences in ANA IIF results, predefined performance criteria are necessary. A within- and between-run reproducibility experiment was performed to reveal imprecision results for the different sample iQC materials used. The within-run precision is estimated by measuring 5 replicates of each iQC material in a single run on a single day. For the between-run reproducibility, each iQC material was tested in duplicate during 10 days. Using an Excel file (Microsoft, Redmond, WA, USA), drawn up in accordance with the CLSI EP5-A2 protocol, the standard deviation (SD) and coefficient of variation (CV) of the observed data were calculated (24).

Results

The within-run imprecision for the negative, positive (speckled) and positive (homogenous) sample iQC is shown in Attachment 1.

The results of the between-run experiment are shown in Attachment 2. An outlier test revealed no outliers in the data set (Grubbs' tests; $p<0,05$; Attachment 3).

The total imprecision of the QUANTA-Lyser - NOVA View system, determined in accordance to the CLSI EP5-A2 protocol is shown in Table 2. There was no total imprecision result obtained for the negative kit iQC, as this

material only generated a fluorescence intensity of '0'. The positive speckled and homogeneous sample iQC materials revealed similar CV% of respectively 27,1% and 25,7%.

Table 2. Total imprecision

CLSI EP5-A2 protocol (No outliers with Grubbs' test)					
	Negative kit iQC (LIU)	Positive kit iQC (LIU)	Negative sample iQC (LIU)	Positive sample iQC speckled (LIU)	Positive sample iQC homogeneous (LIU)
Mean	0	2133,4	13,85	189,35	264,80
Total imprecision (SD)	0	658,0	5,15	51,40	68,17
Total imprecision (CV%)	-	**30,8%**	**37,2%**	**27,1%**	**25,7%**

Based on the total imprecision results of the sample iQC materials, we defined a target CV for the iQC of 25%, 1_{S2} (a single control measurement exceeds the mean +/- 2 CV% target) as a warning limit and 1_{S3} (a single control measurement exceeds the mean +/- 3 CV% target) as stop limit. For patient samples, it is commonly accepted that a variation in end-point titer of one dilution is acceptable. A change of titer equal or more than two steps has to be considered as a clinical significant variation. For the positive speckled sample iQC, a decrease of > 3*CV% (75%), results in a LIU below the diagnostic cut-off (LIU 48) and corresponds to a clinically relevant ANA IIF result of > 1 titer (target: 1:160; >3CV%: < 1:80). Although an increase of >3*CV% of the negative sample iQC of the study did not result in a change of > ANA IIF titer, we have decided to evaluate the results in accordance with acceptation criteria for the negative and positive sample iQC. An overview of the target values and acceptance criteria for the positive and negative iQC are listed Table 3, 4 and 5.

Table 3. Target values positive iQC materials

iQC target values									
%CV	Criteria	Positive kit iQC		Positive speckled sample iQC			Positive homogeneous sample iQC		
		SD (Intensity)	Range (Intensity)	SD (Intensity)	Range (Intensity)	SWT	SD (Intensity)	Range (Intensity)	SWT
25%	-	533,35	1600,1 – 2666,8	47,34	142,0 – 236,7	160 – 160	66,2	198,6 – 331,0	160 – 320
50%	Warning rule	1066,7	1066,7 – 3200,1	94,68	94,7 – 284,0	160 – 320	132,4	132,4 – 397,2	160 – 320
75%	Stop rule	1600,05	533,4 – 3733,5	142,01	47,3 – 331,4	0 – 320	198,6	66,2 – 463,4	80 – 320

Table 4. Target values negative sample iQC

iQC target values				
%CV	Criteria	Negative sample iQC		
		SD (Intensity)	Range (Intensity)	SWT
25%	-	3,46	10,4 – 17,3	N.v.t.
50%	Warning rule	6,93	6,93 – 20,8	N.v.t.
75%	Stop rule	10,39	3,47 – 24,2	N.v.t.

The iQC acceptance criteria for quality indicators are added to Table 5. The same target CV of 25% was applied to the percentage of positive ANA IIF patient samples in a run and for the median and mean patient sample LIU of a run.

Table 5. Acceptance criteria for iQC

		Target value and coefficient of variation (CV)	Acceptance criteria
Process control	LIU positive kit iQC	• Pattern of initial ANA IIF analysis • Target LIU-value: >48 • Target CV: 25%	• Exact match of target pattern • Westgard rules (1_{S2} as a warning limit and 1_{S3} as stop limit)
	LIU negative kit iQC	• Negative on 1:80 dilution • Target LIU-value: ≤48 • Target CV: 25%	• Negative • Westgard rules (1_{S2} as a warning limit and 1_{S3} as stop limit)
	LIU positive sample iQC (speckled/homogeneous)	• Pattern of initial ANA IIF analysis • Target LIU-value: +/- 200 • Target CV: 25%	• Exact match of target pattern • Westgard rules (1_{S2} as a warning limit and 1_{S3} as stop limit)
	LIU negative sample iQC (patient pool)	• Negative on 1:80 dilution • Target LIU-value: ≤48 • Target CV: 25%	• Negative • Westgard rules (1_{S2} as a warning limit and 1_{S3} as stop limit)
Monitoring of patient results	% positive ANA IIF patient samples/run	• Target value: positive/negative ratio at 1:80 dilution of a real-life routine run • Target CV: 25%	• Westgard rules (1_{S2} as a warning limit and 1_{S3} as stop limit)
	Median patient sample LIU/run	• Target value: overall median of the 16 study patient samples (distribution of the LIU-values at 1:80 dilution of a real-life routine run) • Target CV: 25%	• Westgard rules (1_{S2} as a warning limit and 1_{S3} as stop limit)
	Mean patient sample LIU/run	• Target value: overall mean of the 16 study patient samples (distribution of the LIU-values at 1:80 dilution of a real-life routine run) • Target CV: 25%	• Westgard rules (1_{S2} as a warning limit and 1_{S3} as stop limit)

2) Evaluation of different quality indicators in an experimental setup.

Materials and methods

Quality materials

All iQC materials as described in section I, were included in part 2 of the study.

Study samples

Next to the iQC material described in section 1, 16 anonymized rest routine ANA IIF samples were included. Those samples were retrospectively selected, to obtain in terms of positive/negative ratio and distribution of the different LIU-values and corresponding SWT, a simulation of a real-life routine ANA IIF run (Attachment 4). For every patient sample, the relative change of LIU in comparison with the LIU obtained in the 'reference run', performed under controlled circumstances, was calculated. The same target CV (25%) as applied for the performance criteria was used to decide whether the relative change was clinically significant (3*CV%) and corresponded with a stop limit.

Errors in ANA IIF

Artificial errors, mimicking plausible errors in routine ANA IIF practice, were included at different stages of the ANA IIF process. Ten different ANA IIF runs were performed, each involving one error. For every ANA IIF run, the different quality indicators of part I were calculated and evaluated against the predefined performance criteria. An overview of the artificial errors included in the study is listed in Table 6.

Table 6. Errors ANA IIF analysis

Errors in the ANA IIF analytical process	
Pre-analytical problems	
1. Needle obstruction	• 5 µL sample volume in 790 µL PBS-buffer, instead of 10 µL + incubation of 18 µL sample dilution/kit iQC and 18 µL conjugate on slide instead of 35 µL
Analytical problems	
1. PBS-buffer dilution	• 1 bottle PBS-buffer in 2000 mL instead of 1 bottle in 1000 mL
2. Old PBS-buffer	• Use of PBS-buffer after 3 months. Insert claims 4 weeks stability after dilution
3. Old conjugate	• Use of conjugate 3 months after opening
4. Contrad dilution	• 8 mL Contrad in 1000 mL instead of 4 mL in 1000 mL
5. Sample wash step error	• 1 wash cycle with 1 mL instead of 3 wash cycles with 2 mL
6. Conjugate wash step error	• 1 wash cycle with 1 mL instead of 3 wash cycles with 2 mL
7. Needle contamination	• Absence of Contrad buffer (no rinsing liquid)
Post-analytical problems	
1. Final slide incubation >3h	• Slide more than 3 hours in PBS-buffer on QUANTA-Lyser before NOVA View analysis
2. Rescanning slide	• 5x rescanning of same slide

Results

An overview of all the results obtained from the different experiments is given in attachment 5.

Pre-analytical problems

- Needle obstruction

During this experiment an obstruction of the needle was simulated by pipetting only 5 µL (instead of 10 µL) sample in 790 µL PBS-buffer and a smaller volume of diluted sample or iQC and conjugate on the slide (Table 6). This error had a manifest influence on all samples. 88% (n=21/24) of the samples exhibited a relative change in FI of more than 2*CV (50%) with 79% (n=19/24) of the samples exhibiting even more than 3*CV (75%). The LIU of the positive sample iQC, both homogeneous and speckled, as well as the LIU-median exceeded the stop limit. For the LIU-results of the positive or negative kit iQC no significant change in LIU was observed.

The LIU of the negative sample iQC and the percentage of ANA IIF positive patients per run exceeded our predefined warning limit. In contrast to the median, the evaluation of the average of the patients LIU-results revealed no problems.

Analytical problems

- PBS-buffer dilution

During this manipulation, one bottle of PBS-buffer was diluted in 2000 mL instead of 1000 mL. The LIU of 75% of the samples changed relatively more than 3*CV compared to the reference-LIU. The LIU of the positive sample iQC, the LIU of the negative sample iQC, the median-LIU of the patient samples, and additionally the mean LIU of the patient samples exceeded the stop limit. Similar to the needle-obstruction experiment, the percentage of positive ANA IIF patients exceeded the warning limit. The evaluation of the LIU values of neither the positive nor the negative kit iQC indicated any quality problem.

- Old PBS-buffer

The insert of PBS-buffer claims stability after dilution for four weeks. In this experiment a buffer was used which was diluted for three months. No significant difference in LIU could be observed for the study samples, sample iQC's, the median or mean LIU or kit iQC's. The percentage of positive ANA IIF patients remained unchanged.

- Old conjugate

Nearly half (46%, n=11/24) of the samples showed a relative change in LIU of >50% when using a conjugate three month after first opening. No stop limit was obtained. Same influence was seen on the LIU of the positive sample iQC and on the median LIU.

- Contrad dilution

A double-concentrated Contrad-solution did not result in any significant LIU-change of more than 75% compared to the reference run. None of the quality-indicators exceeded a warning or stop limit.

- Sample wash step error

After sample and conjugate incubation, multiple wash cycles are carried out by the QUANTA-Lyser. Changing the quantity and the volume of the wash cycles after sample incubation, did not result in any significant changes of the LIU-results or the evaluation of the quality indicators.

- Conjugate wash step error

In contrast to the previous experiment, lowering the quantity and volume of the wash cycles after conjugate incubation did have an influence on the LIU of the patient samples. The LIU of 58% (n=14/24) of the samples relatively changed more than 2*CV (50%) compared to the reference LIU, with even 25% (n=6/24) of the samples changing more than 3*CV (75%). No influence was noted after the evaluation of the positive or negative kit iQC. Only the median LIU of the patient samples exceeded the stop limit. All other quality indicators remained within warning limits.

- Needle contamination

To simulate needle contamination, no Contrad-buffer was used, resulting in no rinsing liquid and consequently no cleaning of the needle between the pipetting of the samples.
67% (n=16/24) of the samples showed a relative change in LIU of more than 75 % compared to the reference LIU. The LIU of the positive sample iQC, LIU of the negative sample iQC as well as the median LIU of the patient samples exceeded the stop limit. The percentage of positive ANA IIF patients per run exceeded the warning limit. No problems emerged after the evaluation of the negative and positive kit iQC, nor the mean LIU of the patient results.

Post-analytical problems

- Final slide incubation >3h

No significant influence was seen when the stained HEp-2 slide was exposed for more than 3 hours in PBS-buffer on QUANTA-Lyser before NOVA View analysis. None of the predefined quality indicators exceeded a warning or stop limit.

- Rescanning slide

After three, four and five times rescanning, the evaluation of the LIU of respectively 13%, 21% and 53% of the study samples revealed a warning limit (change of 50%), but no stop limit, in concordance with the median LIU. The positive and negative kit iQC showed no significant change in LIU.

3) Evaluation of different quality indicators in daily routine laboratory practice.

Materials and methods

iQC acceptance criteria

For the different quality indicators defined in part 1 (kit iQC materials, the sample iQC materials and the quality markers for the monitoring of patient results) the imprecision data were retrospectively calculated for 10 consecutive, stable routine ANA IIF runs. Only runs using the same lot number of HEp-2 cells, the same sample iQC materials and with a minimum of 20 patient screening (1:80 dilution) samples or more were included in the analysis. The retrospectively obtained 'routine' imprecision data were compared to the predefined 'study' acceptation criteria of part 1.

Quality indicators

The applicability of the selected quality indicators for ANA IIF, defined in part 1, in the daily routine laboratory practice was retrospectively investigated during 3 pre-defined periods of routine ANA IIF analysis in the laboratory of OLV Hospital Aalst in 2017:

- during a stable period (2/08/2017-11/09/2017), without any technical QUANTA-Lyser-NOVA View instrument intervention and using the same HEp2-kit lot number and sample iQC materials
- during two periods (17/05/2017-28/06/2017 and 15/11/2017-31/12/2017) containing a methodological or technical intervention. The presence of the problem was noticed by routine sample iQC violations.

During the three periods, the different quality indicators were calculated and plotted in Levey-Jennings charts. The predefined acceptance criteria (CV 25%) with the corresponding Westgard-rules (1_{S2} as a warning rule and 1_{S3} as stop limit) were used to detect violations.

Results

iQC acceptance criteria

The imprecision data of the retrospective analysis of the different quality indicators are shown in Table 7. For the positive kit iQC, a lower standard deviation was obtained in daily routine practice compared to the experimental setup in part 1. Since the negative kit iQC only generated FI results of '0' or '1', a CV% of more than 200% was revealed. Consequently, the CV% of the negative kit iQC as a quality indicator revealed to be of little significance. The variation of the positive speckled sample iQC (CV 34,7%) exceeded the variation found in the experimental setup (CV 27,1%) and was more in concordance with the CV% of the negative sample iQC, both in the experimental as well as in routine practice. Our retrospective analysis revealed a relatively high variation (CV 47,1%) for the median and mean patient sample LIU per run.

Table 7. Imprecision quality indicator in daily routine

Imprecision quality indicators (10 stable routine runs)							
	LIU Positive kit iQC	LIU Negative kit iQC	LIU Positive sample iQC speckled	LIU Negative sample iQC	% positive ANA IIF patient samples/run	Median patient sample LIU/run	Mean patient sample LIU/run
Mean	2089,2	0,2	269,6	33,3	0,6	79,1	237,58
SD	148,0	0,4	93,5	11,0	0,2	37,3	115,75
CV (%)	**7,1%**	**210,8%**	**34,7%**	**33,1%**	**26,4%**	**47,1%**	**48,72%**

Quality indicators

- Stable period

The Levey-Jennings plots of the different quality indicators in the stable period are given in attachment 6. There were no clinically significant stop limit fluctuations of the predefined quality indicators during this stable period, with the exception of the median LIU and mean LIU, revealing a high variability.

- Unstable period 1

On 07/09/2017, the LIU of the negative and positive sample iQC exceeded the higher 1_{S3} stop limit (Levey-Jennings plots in Attachment 7). After the wash and dilution buffer had been replaced, the results of the rerun on the same day returned to normal. The positive and negative kit IQC remained stable during the period and could not detect the problem present.

- Unstable period 2

During the second unstable period, the negative and positive sample iQC exceeded the lower 1_{S3} stop limit on 06/12/2017 (Attachment 8), confirmed by a 1_{S3} stop violation of the median LIU. No problem could be detected based on the follow-up of the FI of the positive kit iQC, negative kit iQC and mean LIU. The sample needle of the QUANTA Lyser was rinsed with methanol and a replacement of the conjugate was performed, resulting in a normalization of the quality indicators. During the whole unstable period 2, several 'isolated' 1_{S3} stop rule violations of the median LIU per run were found.

Discussion

Two recent Belgian multicenter studies among laboratories performing ANA IIF analysis by NOVA View reported clinically important intra- and inter-laboratory variation in ANA IIF analysis (17, 18). Actually, the large inter-laboratory variation could be observed for every type of automated ANA IIF microscope used in Belgian routine laboratory practice (18). ANA IIF analysis intrinsically bears important analytical variables (e.g. lot number related differences in substrate and conjugate, the subjective result interpretation), that contribute to the large inter-assay variability (7, 16). The primary aim of the implementation of automated ANA IIF microscopes, was a workload reduction on the one hand, but also the harmonization of the ANA IIF analytical process on the other hand (7, 14, 16, 25). Unfortunately, both Belgian multicenter studies showed that even with the automated ANA IIF microscopes, ANA testing in clinical practice remains challenging.

Besides the persistent large inter-assay variability inherent to the analysis, the study of Van den Bremt *et al.* also revealed pre-analytical errors, i.e. problems with the washing unit of the pipetting system, and analytical errors, i.e. calibration, which were not uncovered by the routinely used iQC program, solely based on kit iQC materials (16). The routinely obtained CV% results of the kit iQC materials in part 3 of our study revealed that both the positive kit iQC (low CV%) and negative kit iQC (high CV%) are of little significance as a quality indicator. This was confirmed by the experimental set up as well as the retrospective survey, in which none of the (artificially induced) errors were highlighted by an iQC violation of the FI of the kit iQC materials. To perform an adequate quality assurance of the daily routine ANA IIF by automated instruments, additional quality indicators covering the entire ANA IIF process are necessary. Herein, the selection of adequate control materials is the most critical aspect in getting a thorough quality assurance program. First, it is important that the iQC material assures the whole ANA IIF process, from dilution up to result interpretation, which is not always the case for kit QC. The latter are mostly 'ready for use' and do not require predilution, which routine patient samples do require. Second, the variations in FI of patient-derived iQC materials revealed important attribution to the whole quality assurance process, as titer changes of > 1 correspond to clinically important results (25). From this perspective, positive sample iQC materials with a moderate FI (corresponding to an ANA IIF titer of 1:160) reveal the most useful information.

To evaluate the predefined acceptance criteria and quality indicators of the experimental study in daily practice, routinely derived iQC data were retrospectively investigated. In general, the variation of most quality controls exceeded the variation found in the experimental setup. This finding confirms the persisting high inter-assay variability inherent to the ANA IIF analysis despite automation in the pre-analytical (QUANTA Lyser) and analytical (NOVA View) phase and underlines the need for more efforts in harmonizing automated ANA IIF analysis (18). However, the evaluation of the different quality indicators in the experimental setup revealed that follow-up of LIU values by applying traditional Westgard multi-rules (1_{S2} as a warning limit and 1_{S3} as stop limit) assisted in the analytical and clinical assurance of an ANA IIF run. Nevertheless, the high intrinsic CV% of the ANA IIF analysis does not allow for the application of the 2*CV% (50%) limit for the FI of sample iQC materials as a quality control limit. At the most, this 2*CV% limit can be regarded as a warning signal and to encourage iQC trend analysis. However, a clinical defined 75% stop limit should result in a root cause analysis and a review

of the acceptance of the whole ANA IIF analytical run, definitely if different quality indicators exceed this limit. In the retrospective analysis of the two unstable periods, many of our predefined quality indicators exceeded the 1_{3S}-rule on the day the problem occurred, indicating that the accuracy of the ANA IIF results was no longer guaranteed.

As indicated by Maenhout and colleagues, it is worthwhile to include iQC monitoring based on whole-run ANA IIF patient results in the daily routine iQC analysis (19). The experimental study as well as the retrospective study revealed that the percentage ANA positive samples per run confirmed in most cases the findings of the other quality indicators. Both studies showed that the use of the median LIU of the patient samples per run and not the mean LIU yield the most relevant information. The introduction of an artificial error, which had an impact on patient results, was always accompanied by a violation of the median LIU of the patient samples per run. However, our retrospective analysis revealed a relatively high variability of the median patient sample LIU per run. Several 1_{3S} stop limit violations of the median LIU per run were found over a period of time, without violation of any other quality indicator. Variations in demographic features of patients (e.g. age, gender, hospitalization status, clinical discipline of requester) involved in the ANA IIF analysis run, contribute to this large inter-run variability of the median LIU. A further refinement of median LIU calculations is warranted. Awareness is important so that in daily routine practice decisions are never taken by interpreting only one quality indicator.

Conclusion

The final purpose of this study was to optimize the iQC program of the QUANTA-Lyser-NOVA View system in order to contribute to more accurate ANA IIF reporting and globally to more harmonization in ANA IIF analysis. Based on the results of the experimental study and the retrospective data analysis, we propose the following quality control procedure:

1. Analytical and clinical process control by monitoring the LIU values of the company and patient-derived negative and positive iQC control materials. Patient-derived iQC samples are necessary to ensure that the whole ANA IIF analysis process is controlled, from dilution up to result interpretation.
2. Well-chosen target value of iQC control materials to detect clinically important iQC violations.
3. A target CV for the iQC of 25% can be used, with 1_{S3} as stop limit. The 1_{S2} can be used as warning limit, only to encourage trend analysis or to indicate that further follow-up is required without a clinical problem at that moment.
4. iQC monitoring of patient results based on the percentage ANA positive samples per run and on the median-LIU per run. One run must contain at least 20 patient screening (1:80 dilution) samples in order to calculate these quality indicators.
5. In daily routine practice, decisions can never be taken by interpreting only one quality indicator.

TO DO/ACTIONS

1) Discussion of the proposed quality procedure during a user meeting with the other NOVA View users.
2) Further refinement of the quality procedure in daily practice.
3) Supporting national and international initiatives regarding ANA IIF harmonization.

ATTACHMENTS

Attachment 1: Within-run precision

Within-run precision			
	Negative sample iQC (LIU)	Positive sample iQC speckled (LIU)	Positive sample iQC homogeneous (LIU)
Result 1	15	192	322
Result 2	15	248	318
Result 3	14	217	305
Result 4	17	191	304
Result 5	24	160	250
Mean	17,0	201,6	299,8
SD	4,1	32,9	28,9
CV (%)	**23,9%**	**16,3%**	**9,7%**

Attachment 2: Between-run precision

Between-run precision					
	Negative kit iQC (LIU)	Positive kit iQC (LIU)	Negative sample iQC (LIU)	Positive speckled sample iQC (LIU)	Positive homogeneous sample iQC (ILIU)
Run 1	0	2037	12 11	153 145	214 268
Run 2	0	2120	19 17	255 241	352 344
Run 3	0	2038	20 13	229 188	326 308
Run 4	0	2183	17 17	156 255	317 301
Run 5	0	2057	8 15	151 233	267 273
Run 6	0	2315	17 10	217 196	290 225
Run 7	0	2145	15 8	177 192	207 197
Run 8	0	2316	7 11	112 116	163 259
Run 9	0	1851	25 20	252 261	315 357
Run 10	0	2272	9 6	108 150	216 97

Attachment 3: Outlier check (Grubbs' test) (21)

- NEGATIVE SAMPLE iQC

Descriptive Statistics

Mean: 13.85
SD: 5.15
# of values: 20
Outlier detected? No
Significance level: 0.05 (two-sided)
Critical value of Z: 2.70824545658

Your data

Row	Value	Z	Significant Outlier?
1	12.	0.36	
2	11.	0.55	
3	19.	1.00	
4	17.	0.61	
5	20.	1.19	
6	13.	0.16	
7	17.	0.61	
8	17.	0.61	
9	8.	1.14	
10	15.	0.22	
11	17.	0.61	
12	10.	0.75	
13	15.	0.22	
14	8.	1.14	
15	7.	1.33	
16	11.	0.55	
17	25.	2.16	Furthest from the rest, but not a significant outlier (P > 0.05).
18	20.	1.19	
19	9.	0.94	
20	6.	1.52	

- POSITIVE SAMPLE iQC (SPECKLED)

Descriptive Statistics

Mean: 264.80
SD: 68.17
# of values: 20
Outlier detected? No
Significance level: 0.05 (two-sided)
Critical value of Z: 2.70824545658

Your data

Row	Value	Z	Significant Outlier?
1	214.	0.75	
2	268.	0.05	
3	352.	1.28	
4	344.	1.16	
5	326.	0.90	
6	308.	0.63	
7	317.	0.77	
8	301.	0.53	
9	267.	0.03	
10	273.	0.12	
11	290.	0.37	
12	225.	0.58	
13	207.	0.85	
14	197.	0.99	
15	163.	1.49	
16	259.	0.09	
17	315.	0.74	
18	357.	1.35	
19	216.	0.72	
20	97.	2.46	Furthest from the rest, but not a significant outlier (P > 0.05).

- POSITIVE SAMPLE iQC (HOMOGENEOUS)

Descriptive Statistics

Mean: 189.35
SD: 51.40
# of values: 20
Outlier detected? No
Significance level: 0.05 (two-sided)
Critical value of Z: 2.70824545658

Your data

Row	Value	Z	Significant Outlier?
1	153.	0.71	
2	145.	0.86	
3	255.	1.28	
4	241.	1.00	
5	229.	0.77	
6	188.	0.03	
7	156.	0.65	
8	255.	1.28	
9	151.	0.75	
10	233.	0.85	
11	217.	0.54	
12	196.	0.13	
13	177.	0.24	
14	192.	0.05	
15	112.	1.50	
16	116.	1.43	
17	252.	1.22	
18	261.	1.39	
19	108.	1.58	Furthest from the rest, but not a significant outlier (P > 0.05).
20	150.	0.77	

Attachment 4: Overview of the ANA IIF characteristics of the patient samples included in the study

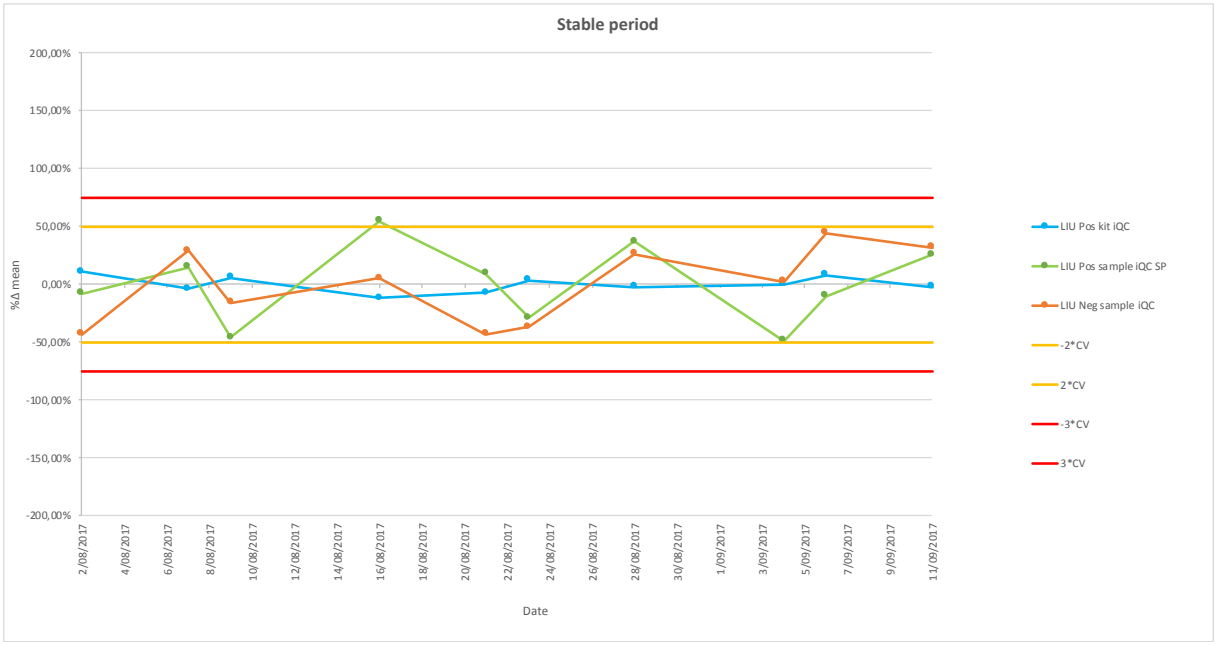
Patient samples			
Study number	**LIU**	**ANA IIF Pattern**	**SWT**
Patient sample 1	36	-	-
Patient sample 2	24	-	-
Patient sample 3	32	-	-
Patient sample 4	25	-	-
Patient sample 5	33	-	-
Patient sample 6	10	-	-
Patient sample 7	9	-	-
Patient sample 8	55	Homogeneous	80
Patient sample 9	82	Speckled	80
Patient sample 10	120	Speckled	80
Patient sample 11	188	Speckled	160
Patient sample 12	213	Homogeneous	160
Patient sample 13	443	Homogeneous	320
Multicenter study sample 2	434	Speckled	320
Patient sample 14	1050	Homogeneous	640
Patient sample 15	2228	Speckled	1280

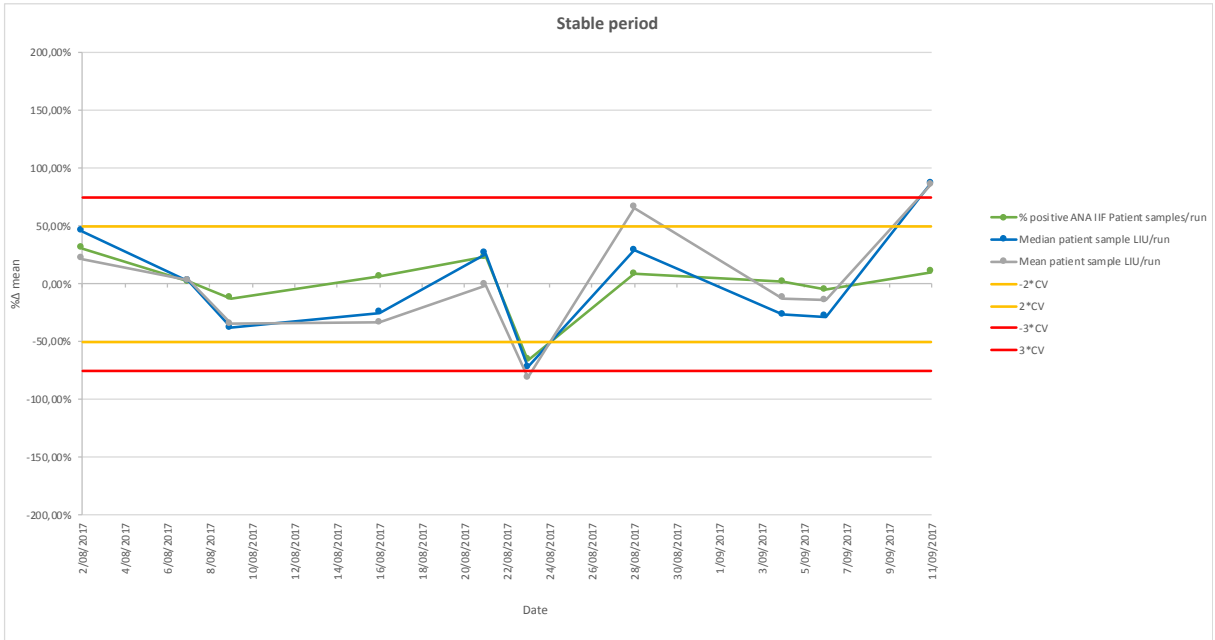
Attachment 5: Effect of errors in ANA IIF

Effect of errors on iQC performance							
	Rescanning 5x				Slide incubation >3h PBS	Old conjugate 3m	Needle contamination
	Scan 2	Scan 3	Scan 4	Scan 5			
LIU Pos kit iQC	-11,08%	-13,25%	-19,72%	-26,32%	-9,43%	-1,84%	-0,77%
LIU Neg kit iQC	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
LIU Pos sample iQC SP	-35,08%	-45,36%	-50,20%	-69,35%	-40,52%	-49,19%	194,74%
LIU Pos sample iQC HOM	-17,67%	-49,57%	-39,51%	-63,22%	-25,57%	-56,61%	134,85%
LIU Neg sample iQC	-16,67%	-11,11%	-33,33%	-47,22%	-11,11%	-44,44%	1275,99%
% positive ANA IIF Patient samples/run	-8,57%	-4,00%	-10,00%	-10,00%	0,00%	-10,00%	50,00%
Median patient sample LIU/run	-14,94%	-41,46%	-54,88%	-65,85%	-24,70%	-60,06%	166,60%
Mean patient sample LIU/run	-1,13%	-16,23%	-30,05%	-38,28%	-14,15%	-23,74%	46,19%
Δ titer step (≥ 2 steps)	0	0	0	0	0	0	2

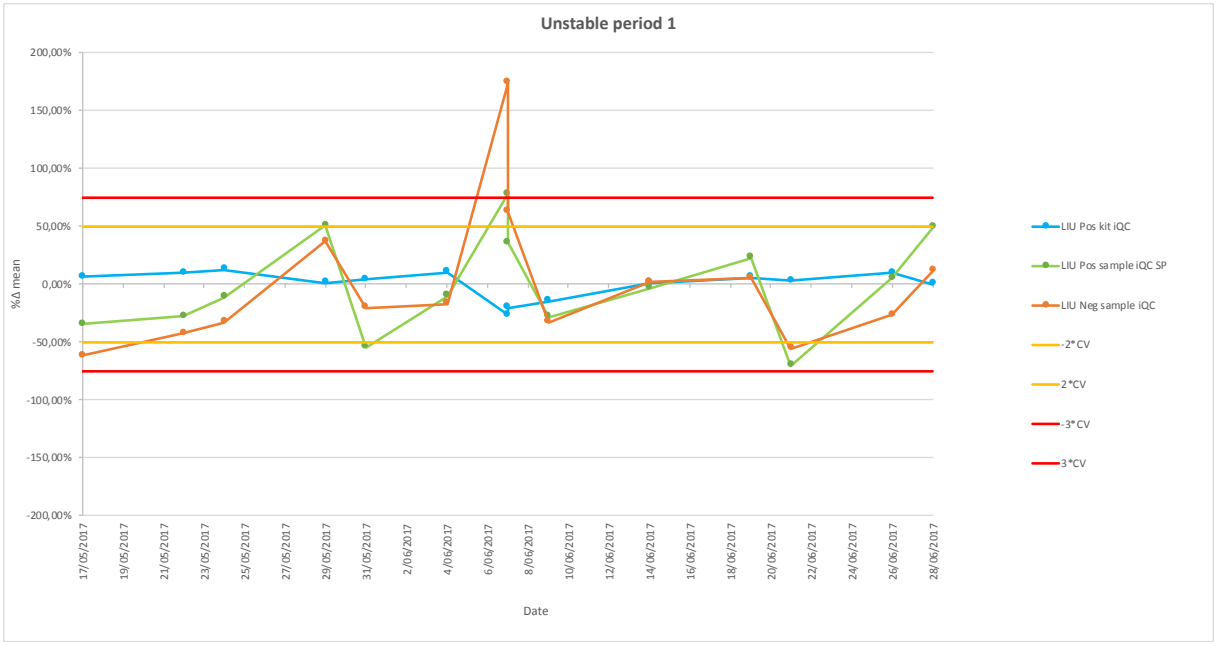
Effect of errors on iQC performance						
	Needle obstruction	Contrad dilution	Buffer dilution	Sample wash step	Conjugate wash step	Old buffer
LIU Pos kit iQC	1,35%	6,18%	-2,41%	-10,47%	-2,47%	-4,16%
LIU Neg kit iQC	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
LIU Pos sample iQC SP	-85,65%	-30,44%	105,85%	-43,75%	43,06%	-8,13%
LIU Pos sample iQC HOM	-92,19%	-33,05%	123,56%	-39,22%	64,60%	-12,78%
LIU Neg sample iQC	-100,00%	-25,00%	1430,56%	-22,22%	1,93%	1,93%
% positive ANA IIF Patient samples/run	-62,50%	0,00%	60,00%	10,00%	25,00%	0,00%
Median patient sample LIU/run	-85,97%	-33,23%	303,35%	-37,50%	89,10%	-9,13%
Mean patient sample LIU/run	-43,25%	-13,75%	92,58%	-8,33%	22,67%	3,40%
Δ titer step (≥ 2 steps)	9	0	9	0	0	0

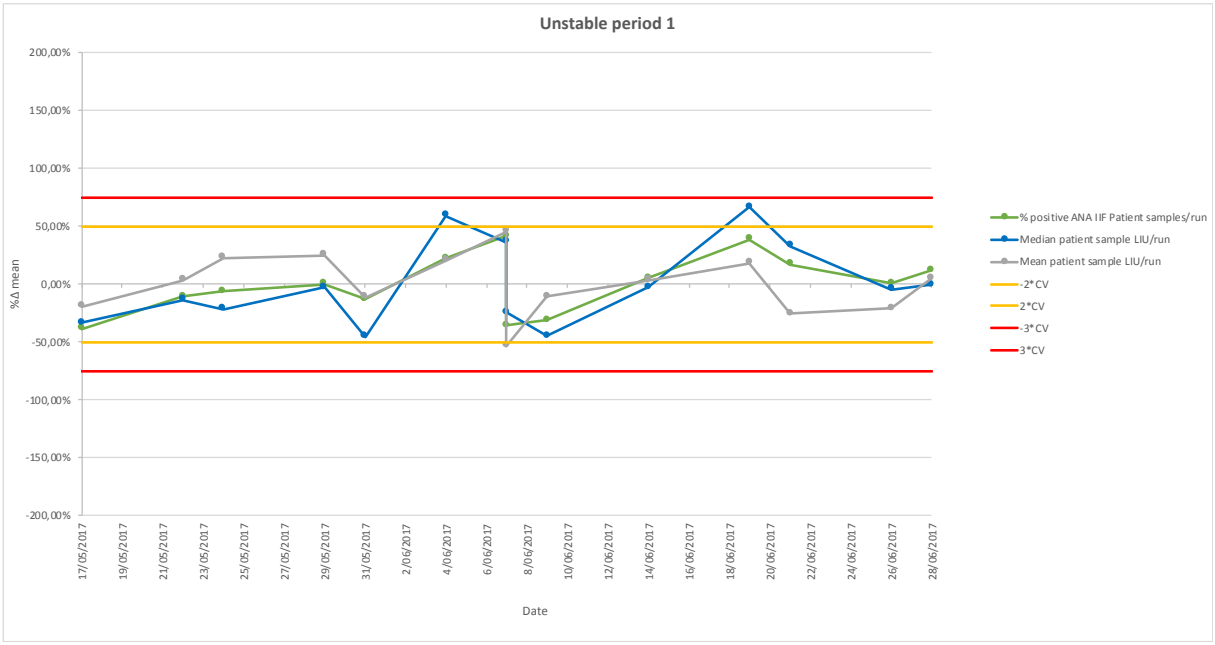
Attachment 6: Evaluation of different quality indicators in daily routine laboratory practice (stable period)



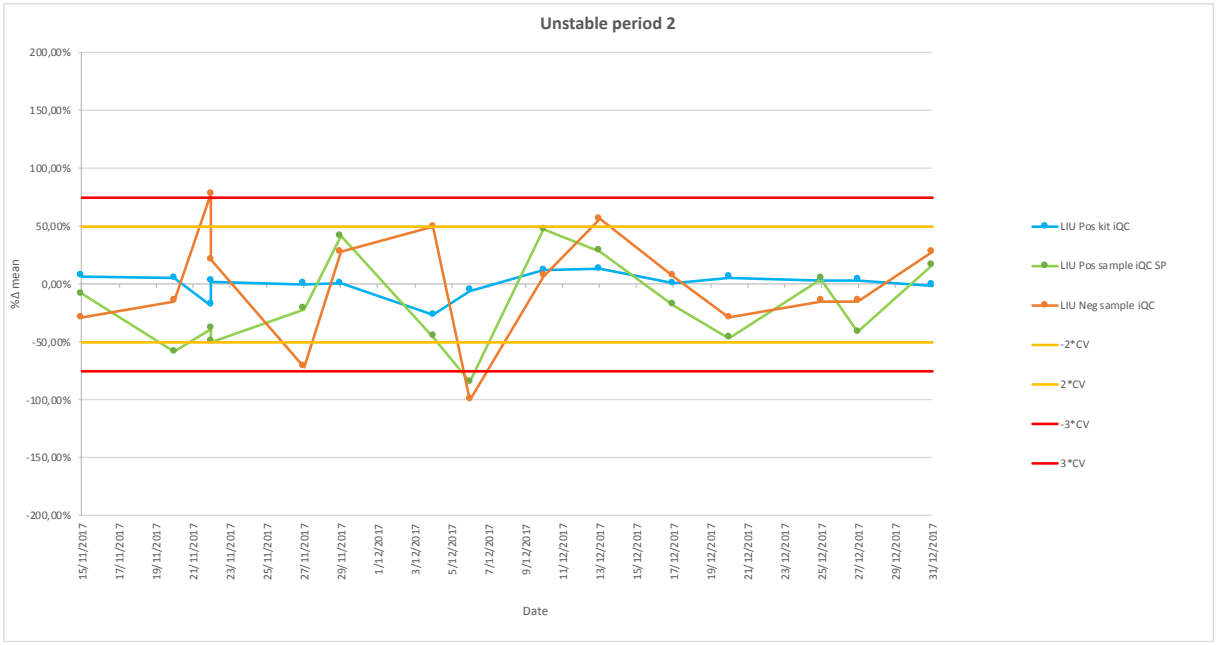


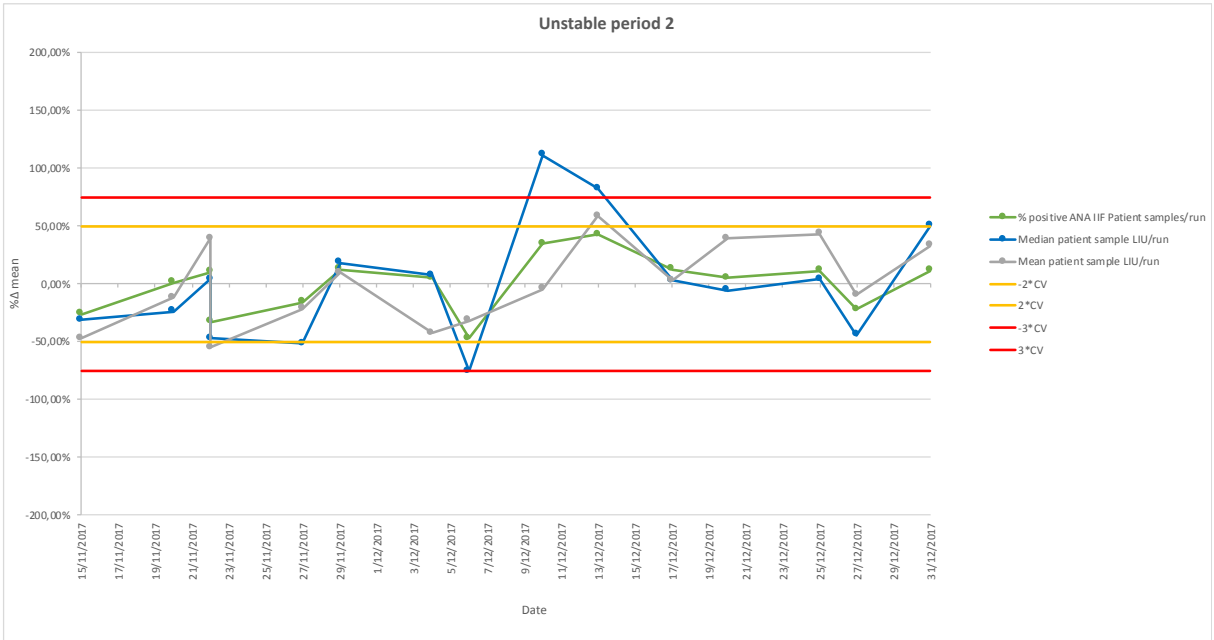
Attachment 7: Evaluation of different quality indicators in daily routine laboratory practice (unstable period 1)





Attachment 8: Evaluation of different quality indicators in daily routine laboratory practice (unstable period 2)





RELEVANT EVIDENCE/REFERENCES

1. Solomon DH, Kavanaugh AJ, Schur PH, American College of Rheumatology Ad Hoc Committee on Immunologic Testing Guidelines. Evidence-based guidelines for the use of immunologic tests: antinuclear antibody testing. Arthritis & Rheumatology. 2002;47(4):434-44.
2. Fritzler MJ. Choosing wisely: Review and commentary on anti-nuclear antibody (ANA) testing. Autoimmunity Reviews. 2016;15(3):272-80.
3. Meroni PL, Schur PH. ANA screening: an old test with new recommendations. Annals of Rheumatic Diseases. 2010;69(8):1420-2.
4. Wiik AS, Høier-Madsen M, Forslid J, Charles P, Meyrowitsch J. Antinuclear antibodies: a contemporary nomenclature using HEp-2 cells. Journal of Autoimmunity. 2010;35(3):276-90.
5. Rigon A, Infantino M, Merone M, Iannello G, Tincani A, Cavazzana I, et al. The inter-observer reading variability in anti-nuclear antibodies indirect (ANA) immunofluorescence test: A multicenter evaluation and a review of the literature. Autoimmunity Reviews. 2017;16(12):1224-9.
6. Copple SS, Giles SR, Jaskowski TD, Gardiner AE, Wilson AM, Hill HR. Screening for IgG antinuclear autoantibodies by HEp-2 indirect fluorescent antibody assays and the need for standardization. American Journal of Clinical Pathology. 2012;137(5):825-30.
7. Infantino M, Meacci F, Grossi V, Manfredi M, Benucci M, Merone M, Soda P. The burden of the variability introduced by the HEp-2 assay kit and the CAD system in ANA indirect immunofluorescence test. Journal of Immunology Research. 2016;65(1):345-54.
8. Mahler M, Meroni PL, Bossuyt X, Fritzler MJ. Current concepts and future directions for the assessment of autoantibodies to cellular antigens referred to as anti-nuclear antibodies. Journal of Immunology Research. 2014;2014(315179).
9. Tozzoli R, Bonaguri C, Melegari A, Antico A, Bassetti D, Bizzaro N. Current state of diagnostic technologies in the autoimmunology laboratory. Clinical Chemistry and Laboratory Medicine. 2013;51(1):129-38.
10. Kivity S, Gilburd B, Agmon-Levin N, Carrasco MG, Tzafrir Y, Sofer Y, et al. A novel automated indirect immunofluorescence autoantibody evaluation. Clinical Rheumatology. 2012;31(3):503-9.
11. Krause C, Ens K, Fechner K, Voigt J, Fraune J, Rohwäder E, et al. EUROPattern Suite technology for computer-aided immunofluorescence microscopy in autoantibody diagnostics. Lupus. 2015;24(4):516-29.
12. Bossuyt X, Cooreman S, De Baere H, Verschueren P, Westhovens R, Blockmans D, et al. Detection of antinuclear antibodies by automated indirect immunofluorescence analysis. Clinica Chimica Acta. 2013;415:101-6.
13. Bonroy C, Verfaillie C, Smith V, Persijn L, De Witte E, De Keyser F, et al. Automated indirect immunofluorescence antinuclear antibody analysis is a standardized alternative for visual microscope interpretation. Clinical Chemistry and Laboratory Medicine. 2013;51(9):1771-9.
14. Meroni PL, Bizzaro N, Cavazzana I, Borghi MO, Tincani A. Automated tests of ANA immunofluorescence as throughput autoantibody detection technology: strengths and limitations. BMC Medicine. 2014;12:38.
15. Bizzaro N, Antico A, Platzgummer S, Tonutti E, Bassetti D, Pesente F, et al. Automated antinuclear immunofluorescence antibody screening: a comparative study of six computer-aided diagnostic systems. Autoimmunity Reviews. 2014;13(3):292-8.

16. Mahler M. Lack of standardisation of ANA and implications for drug development and precision medicine. Annals of the Rheumatic Diseases. 2018; doi: 10.1136/annrheumdis-2018-213374.
17. Van den Bremt S, Schouwers S, Van Blerk M, Van Hoovels L. ANA IIF Automation: Moving towards Harmonization? Results of a Multicenter Study. Journal of Immunology Research. 2017;2017:6038137.
18. Van Hoovels L, Schouwers S, Van den Bremt S, Bossuyt X. Variation in antinuclear antibody detection by automated indirect immunofluorescence analysis. Annals of the Rheumatic Diseases. 2018; doi: 10.1136/annrheumdis-2018-213543.
19. Maenhout TM, Bonroy C, Verfaillie C, Stove V, Devreese K. Automated indirect immunofluorescence microscopy enables the implementation of a quantitative internal quality control system for anti-nuclear antibody analysis. Clinical Chemistry and Laboratory Medicine. 2014;52(7):989-98.
20. Mulliez SM, Maenhout TM, Bonroy C. Impact of the routine implementation of automated indirect immunofluorescence antinuclear antibody analysis: 1 year of experience. Clinical Chemistry and Laboratory Medicine. 2016;54(7):e183-6.
21. Schouwers S, Bonnet M, Verschueren P, Westhovens R, Blockmans D, Mariën G, et al. Value-added reporting of antinuclear antibody testing by automated indirect immunofluorescence analysis. Clinical Chemistry and Laboratory Medicine. 2014;52(4):547-51.
22. Clinical and Laboratory Standards Institute (CLSI). Quality assurance of laboratory tests for autoantibodies to nuclear antigens: (1) indirect fluorescent assay for microscopy and (2) microtiter enzyme immunoassay methods. Approved guideline; second edition. Wayne, PA: CLSI; 2006. CLSI document. 2006;I/LA2-A2.
23. Algeciras-Schimnich A, Algeciras-Schimnich A, Bruns DE, Boyd JC, Bryant SC, La Fortune KA, et al. Failure of current laboratory protocols to detect lot-to-lot reagent differences: findings and possible solutions. Clinical Chemistry. 2013;59(8):1187-94.
24. Clinical and Laboratory Standards Institute (CLSI). Evaluation of precision performance of quantitative measurement methods; Approved guideline-Second edition. Wayne, PA: CLSI; 2004. CLSI document. 2004;EP5-A2.
25. Tozzoli R, Villalta D, Bizzaro N. Challenges in the Standardization of Autoantibody Testing: a Comprehensive Review. Clinical Reviews in Allergy and Immunology. 2017;53(1):68-77.